

**OUCH!**

La lettre d'information mensuelle sur la sensibilisation à la sécurité pour vous

Réfléchissez avant d'envoyer votre invite : Utiliser l'IA de manière sécurisée

La chance d'une vie, volée

Léna avait récemment commencé à utiliser un chatbot IA pour l'aider à gérer son emploi du temps. Entre la famille, les factures, et les plans pour le futur, elle adorait à quel point il était facile de poser des questions et d'avoir la réponse instantanément. Un soir, elle s'inquiétait pour ses finances, alors elle a demandé à l'IA de l'aider à investir. L'IA lui a rapidement suggéré une stratégie. Cependant, elle n'avait pas tout le contexte de la situation de Léna, sa tolérance au risque ou les implications fiscales propres à ses circonstances. L'IA lui a recommandé d'investir son argent dans un panier composé d'actions à la mode et d'opérations à court terme qui promettaient de meilleurs rendements. Il lui a même expliqué comment gérer la fiscalité liée à ses investissements.

Le conseil semblait sûr et bien pensé, alors Léna l'a suivi. Au début, elle était excitée. Mais en l'espace de quelques mois, le marché a basculé et ses investissements ont perdu de la valeur. Pire encore, à l'approche de la période des déclarations d'impôts, elle s'est rendu compte qu'elle avait mal compris certaines règles essentielles. Comme elle avait suivi les conseils de l'IA sans les vérifier, elle a commis des erreurs fiscales qui lui ont valu des pénalités et des frais supplémentaires.

Finalement, Léna a perdu de l'argent en suivant des conseils qui ne correspondaient pas à sa situation. L'IA peut être un outil utile, mais il est important de rappeler qu'elle peut faire des erreurs. Lorsque vous lui faites confiance sans vérifier ce qu'elle dit, de petites erreurs peuvent rapidement coûter cher.

Qu'est-ce que l'intelligence artificielle (IA) ?

L'intelligence artificielle (IA) est une technologie conçue pour simuler la manière dont les humains pensent, traitent l'information et prennent des décisions. Cela peut inclure la génération de texte, la reconnaissance d'images, la prise de décisions, la création de contenu ou la résolution de problèmes. En général, il y a trois sortes d'IA que vous pouvez utiliser.

- **L'IA intégrée** : c'est une IA intégrée aux outils que vous utilisez quotidiennement. Souvent, vous utilisez l'IA sans vous en rendre compte. Par exemple, lorsque vous prenez une photo avec votre téléphone, il y a des chances que l'IA améliore votre photo.
- **L'IA générative** : ce sont des services d'IA spécialement conçus pour aider les gens. Ce sont notamment des outils tels que ChatGPT, Google Gemini, ou Anthropic Claude. L'IA générative peut vous aider dans de nombreuses tâches, comme créer de la musique, rédiger un business plan, générer des images ou analyser vos idées.
- **L'IA agentielle** : ce sont des services d'IA conçus pour agir en votre nom. Ces systèmes peuvent fonctionner avec une intervention humaine limitée, voire nulle, et s'intégrer à une main-d'œuvre numérique, en prenant des décisions et en agissant sur la base de lignes directrices ou d'instructions générales.

Utiliser l'IA de manière sécurisée

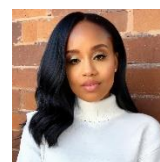
Sur les trois types d'IA, nous allons nous concentrer sur l'IA générative, puisque c'est le type d'IA que vous utilisez le plus. L'IA générative est un outil puissant qui peut vous aider à accomplir vos tâches plus efficacement et à acquérir de nouvelles compétences et idées, à condition de bien l'utiliser. Voici quelques points à considérer :

- **La confidentialité** : méfiez-vous de ce que vous partagez aux outils d'IA. Lorsque vous téléchargez ou saisissez des informations dans un système d'IA, celles-ci peuvent être traitées et stockées, et, dans certains cas, utilisées pour améliorer le service, selon la plateforme. Si vous communiquez des informations très sensibles à une IA, celles-ci pourraient être transmises à des tiers. Ne partagez que ce que vous pouvez dire en toute confiance. Une autre option consiste à recourir à des services d'IA qui protègent votre vie privée en utilisant des modèles qui n'apprennent pas à partir de vos données.
- **L'exactitude** : l'IA peut affirmer que ce qu'elle crée ou partage avec vous est exact, même lorsque ce n'est pas le cas. Vérifiez toujours les informations que vous donne l'IA. Un problème commun avec l'IA est qu'elle essaiera toujours de vous donner une réponse, même lorsqu'elle n'a pas compris votre question. Par défaut, l'IA ne vous demandera pas de clarifier votre question, donc il est important d'être aussi clair que possible et de rester vigilant face à des résultats vagues dans la réponse de l'IA.
- **Les biais** : tout comme les humains qui ont programmé les IA, les IA peuvent avoir leurs propres biais. Cela peut donner lieu à des réponses qui semblent assurées, mais qui ne sont pas forcément équilibrées ou exactes. Un biais très commun de l'IA est que cette dernière cherche tout le temps à vous rendre heureux, donc elle vous dira toujours ce que vous voulez entendre. De plus, l'IA ne "sait" que les informations qu'on lui a apprises et auxquelles elle a accès. Si celles-ci sont fondamentalement fausses ou limitées, ses réponses refléteront les lacunes de ces données.

L'IA est un des outils les plus puissants disponible aujourd'hui. Elle peut vous aider à travailler plus vite, à apprendre davantage de choses et à être plus productif. Mais comme n'importe quel outil puissant, elle doit être utilisée avec précaution. Ne lui faites pas confiance aveuglément. Ne lui partagez pas trop d'informations personnelles. Ne lui donnez pas plus de contrôle que nécessaire. Utilisez l'IA comme un outil pour vous aider dans vos décisions, pas pour remplacer votre jugement.

Rédacteur invité

Portia Jefferson est une experte en cybersécurité spécialisée dans la sécurité de l'IA, la sensibilisation aux risques et les conseils pratiques destinés aux utilisateurs lambda. Forte d'une expérience dans les domaines de la fintech et de la fiscalité, elle aide les particuliers à s'y retrouver en toute sécurité dans les technologies émergentes, tant au travail qu'à la maison.



Ressources

Méfiez-vous des deepfakes, une nouvelle ère de tromperie : <https://www.sans.org/newsletters/ouch/beware-deepfakes-new-age-of-deception>

Les voix fantômes : se défendre contre les attaques par clonage vocal : <https://www.sans.org/newsletters/ouch/phantom-voices-defend-against-voice-cloning-attacks>

Se protéger quand la véritable confidentialité est impossible : <https://www.sans.org/newsletters/ouch/protecting-yourself-when-true-privacy-is-impossible>

Traduit pour la communauté par : Juliette Busson

OUCH ! Publié par SANS Security Awareness et distribué sous licence Creative Commons BY-NC-ND 4.0. Vous êtes libre de partager ou de distribuer cette newsletter à condition de ne pas la vendre ni la modifier. Comité de rédaction : Phil Hoffman, Leslie Ridout, Princess Young.

Vous pouvez trouver plus de contenu OUCH! Sur le lien suivant : <https://www.sans.org/newsletters/ouch>